

证 明

本证明之附件是向本局提交的下列专利申请副本

申 请 日： 2003.12.19

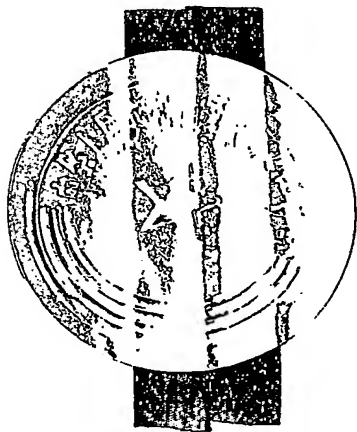
申 请 号： 2003101239793

申 请 类 别： 发明

发明创造名称： 一种多I S P 局域网的出口路由方法

申 请 人： 华为技术有限公司

发明人或设计人： 陈阵、黄河



中华人民共和国
国家知识产权局局长

王 景 川

2005 年 1 月 11 日

BEST AVAILABLE COPY

权 利 要 求 书

1、一种多 ISP 局域网的出口选择方法，包括如下步骤：

将局域网路由设置为普通路由和策略路由，所述的普通路由作为
5 所述的策略路由的备份；

设置所述策略路由的路由策略；

建立 NAT 地址池；

当有来自局域网内部的出口请求时，

查询路由表，确定下一跳的所述策略路由和/或普通路由；

10 判断所述的策略路由是否可用，

如果可用，

用所述的策略路由结果覆盖目的地址路由；

如果不可用，

仍使用原来的目的地址路由；

15 判断是否需要 NAT 转换，

如果需要，

选择所述的 NAT 地址池和出接口用户板；

转分布式 NAT 处理设备，实现 NAT 转换；

按上述的路由结果转发报文到出接口用户板；

20 如果不需要，

根据路由信息转发报文到用户板。

2. 如权利要求 1 所述的多 ISP 局域网的出口选择方法，其特征在于所述的建立 NAT 地址池的步骤包括将所述的 NAT 地址池与各个 ISP 的出接口绑定，和以所述的出接口和访问请求的源 IP 地址的组合为关键字，

创建 NAT 策略树的步骤，其中，所创建的策略树用于存储用户流量是否需要 NAT 转换，所述组合与所述地址池的绑定关系以及分布式专用 NAT 转换板槽号的 NAT 等策略 NAT 消息。

3. 如权利要求 2 所述的多 ISP 局域网的出口选择方法，其特征在于所
5 创建的 NAT 策略树的是二叉树。

4. 如权利要求 2 所述的多 ISP 局域网的出口选择方法，其特征在于所
述的判断是否需要进行 NAT 转换的步骤包括判断用户流量命中的路由
表项中是否有公网标记的步骤，和如果所述的用户流量命中的路由表
项中有公网标记，根据所述的出接口和源 IP 地址的组合为关键字，查
10 询是否命中所述策略树叶子节点。

5. 如权利要求 2 所述的多 ISP 局域网的出口选择方法，其特征在于所
述的选择 NAT 的地址池的步骤包括以所述的出接口和源 IP 地址的组合
为关键字，在所述策略树叶子节点中匹配，在匹配到的 NAT 策略树的
叶子节点得到地址池和 NAT 板号的步骤。

15 6. 如权利要求 1 所述的多 ISP 局域网的出口选择方法，其特征在于所
述的查询路由表，确定下一跳的策略路由和/或普通路由的步骤包括当
所述的策略路由不可用时，用户流量自动按照所述的普通路由转发。

7. 如权利要求 1 所述的多 ISP 局域网的出口选择方法，其特征在于所
述的判断策略路由是否可用的步骤是：

20 用所述的策略路由的下一跳查找所述的路由表；
判断下一跳能够命中 32 位掩码的直连主机路由，
如果是，
所述的策略路由可用；
如果不是，

所述的策略路由不可用，以所述的普通路由转发。

8. 如权利要求 1 所述的多 ISP 局域网的出口选择方法，其特征在于所述的方法中所述的路由模块与 NAT 模块完全分离，所述的路由模块确定用户流量的出口，所述的 NAT 模块确定是否要做 NAT 以及地址池的选择。
9. 如权利要求 1 所述的多 ISP 局域网的出口选择方法，其特征在于所述的方法还包括判断是否存在多个下一跳的步骤，并且当存在多个所述的下一跳时，包括步骤：多个 ISP 分别分担部分流量。

10

15

说明书

一种多 ISP 局域网的出口路由方法

5 技术领域

本发明涉及网络路由技术,具体涉及一种多 ISP 局域网的出口选择方法。更具体涉及一种连接多个 Internet 服务提供商 (ISP) 的局域网的访问出口的选择方法。

10 背景技术

为了便于描绘本发明和背景技术,首先给出说明书和权利要求中使用的短语的定义:

NAT (Network Address Translation): 网络地址转换

ISP (Internet Service Provider): Internet 网络接入服务提供商

15 主机路由: 与网络设备直联主机在路由表中对应的 32 位掩码表项,此类表项都有 ARP 与其一一对应。

20 一般情况下,园区网通常配置多个网络出口,联接ISP(Internet Service Provider)接入服务商,多个网络出口之间互为备份、负荷分担,提高园区网与外网间通信的带宽和鲁棒性。在校园网的建设中,这种情况更为常见,通常校园网会接入到一个公共运营商和教育网。同时,由于目前IP地址非常短缺,园区网中需要使用私网 Ip 地址,通过 NAT (网络地址转换) 接入到 Internet 中。由于不同的接入服务商提供的接入计费、流控策略不同,因此需要对园区网出口流量进行 NAT 多出口策略控制。通过 NAT 多出口策略控制,可以根据用户报文的源信息

和目的信息，来选择出口 ISP，达到节约接入费用，用户分级管理的目的。例如在校园网中，接入公共运营商（如电信）时到境外的流量收费比接入教育网便宜，而教育网对国内流量不收费。因此需要进行 NAT 多出口策略控制，对到国内的流量选择教育网出口，国外的流量选择公共运营商出口。同时两个出口还应该互相备份，当一个出口故障时，所有流量切换另一个出口。

对于 NAT 多出口策略控制，业界主流核心路由交换机对这样的组网没有很好的解决方案，或者束手无策。一般业界采用如下方案来解决 NAT 多出口策略问题。使用多台路由器来组网实现，前面一台路由器进行报文分流，对报文根据源信息和目的信息进行分流，转到后面的多台路由器上，由后面的多台路由器进行 NAT 操作，达到实现 NAT 多出口策略控制的目的；

考虑到园区网联接多个出口 ISP，内部公私网混合组网，多个地址空间并存，园区核心路由交换机要求 NAT 和策略路由并存等多种组网要求，需要根据源 IP 地址和出接口共同决定报文转发是否需要 NAT 转换及其地址池绑定关系，为了方便可以以 NAT 策略树概念来描述地址池绑定关系；同时将策略路由模块与 NAT 模块分离，以适应分布式转发的需求，使核心交换机能单独提供 NAT 多出口策略功能，实现硬件分布式转发，达到线速无阻塞的转发的要求。

图 1 为现有技术的多 ISP 园区网组网的原理图。它使用多台路由器交换机，或者一台核心交换机多台专用 NAT 设备，来实现 NAT 多出口策略功能。这种组网虽然可以实现 NAT 多出口策略功能，同时由硬件进行 NAT 操作，能保证带宽。但是需要添加设备来实现，组网成本增加的同时也导致故障点增加。另外由于是多个设备混和组网实现，

因此当一个 ISP 出口的 NAT 发生故障时，核心路由交换机不能感知该设备故障，将导致出口流量中断，需要手工干预修改复杂流分类策略，来实现多个 ISP 出口的备份。

5 发明内容

本发明的目的是克服现有技术的上述缺点，提供一种多 ISP 局域网的出口选择方法。利用本发明，NAT 的多 ISP 的控制一台核心路由交换机完成处理，从而克服了多台设备带来的成本和故障点增加的问题，并能适应分布式转发的需求，在实现复杂的转发策略的同时，实现线
10 速无阻塞转发，同时能实现出口链路的备份功能。

本发明提供了一种多 ISP 局域网的出口选择方法，包括如下步骤：

将局域网路由设置为普通路由和策略路由，所述的普通路由作为所述的策略路由的备份；

设置所述策略路由的路由策略；

15 建立 NAT 地址池；

当有来自局域网内部的出口请求时，

查询路由表，确定下一跳的所述策略路由和/或普通路由；

判断所述的策略路由是否可用，

如果可用，

20 用所述的策略路由结果覆盖目的地址路由；

如果不可用，

仍使用原来的目的地址路由；

判断是否需要进行 NAT 转换，

如果需要，

选择所述的 NAT 地址池和出接口用户板;
转分布式 NAT 处理设备, 实现 NAT 转换;
按上述的路由结果转发报文到出接口用户板;
如果不需要,

5 根据路由信息转发报文到用户板。

可选地, 所述的建立 NAT 地址池的步骤包括将所述的 NAT 地址池与各个 ISP 的出接口绑定, 和以所述的出接口和访问请求的源 IP 地址的组合为关键字, 创建 NAT 策略树的步骤, 其中, 所创建的策略树用于存储用户流量是否需要 NAT 转换, 所述组合与所述地址池的绑定关系以及分布式专用 NAT 转换板槽号的 NAT 等策略 NAT 消息。

优选地, 所创建的 NAT 策略树的是二叉树。

可选地, 所述的判断是否需要进行 NAT 转换的步骤包括判断用户流量命中的路由表项中是否有公网标记的步骤, 和如果所述的流量命中的路由表项中有公网标记, 根据所述的出接口和源 IP 地址的组合为关键字, 查询是否命中所述策略树叶子节点。

优选地, 所述的选择 NAT 的地址池的步骤包括以所述的出接口和源 IP 地址的组合为关键字, 在所述策略树叶子节点中匹配, 在匹配到的 NAT 策略树的叶子节点得到地址池和 NAT 板号的步骤。

可选地, 所述的查询路由表, 确定下一跳的策略路由和/或普通路由的步骤包括当所述的策略路由不可用时, 用户流量自动按照所述的普通路由转发。

优选地, 所述的判断策略路由是否可用的步骤是:

用所述的策略路由的下一跳查找所述的路由表;

判断下一跳能够命中 32 位掩码的直连主机路由,

如果是,

所述的策略路由可用;

如果不是,

所述的策略路由不可用, 以所述的普通路由转发。

- 5 可选地, 所述的方法中所述的路由模块与 NAT 模块完全分离, 所述的路由模块确定用户流量的出口, 所述的 NAT 模块确定是否要做 NAT 以及地址池的选择。

- 优选地, 所述的方法还包括判断是否存在多个下一跳的步骤, 并且当存在多个所述的下一跳时, 包括步骤: 多个 ISP 分别分担部分流量。
- 10

- 本发明通过根据源 IP 地址、出接口、普通路由、策略路由来共同决定报文转发是否需要 NAT 转换及其地址池绑定关系, 增加 NAT 策略树, 描述地址池绑定关系; 同时路由模块与 NAT 模块分离, 以适应分布
- 15 式转发的需求, 在实现复杂的转发策略的同时, 达到线速无阻塞的转发要求。

附图说明

- 图 1 为现有技术的多 ISP 园区网组网的原理图;
- 20 图 2 示出了本发明的实施例的 NAT 多 ISP 策略转发流程图;
- 图 3 描绘了本发明的另一实施例的结合具体设备描绘的策略 NAT 转发流程图。

具体实施方式

5 了网络的复杂性。

通过增加NAT策略表，完成NAT策略功能，可以解决两个主要技术问题：

- 1、实现了根据用户源信息和出接口信息来选择ISP出口，可以灵活的配合多个ISP的计费流控策略，节省用户的出口流量费用。
- 10 2、实现了多个出口ISP之间的热备份，发现某个ISP出口故障时，可以自动快速切换到另外的ISP出口，无需人工干预。

图2示出了本发明的NAT多ISP策略转发流程图。其中，在转发平面增加一个NAT策略表，策略表的存储方式是树结构。该表的索引是源IP+出接口，表项中内容是绑定的ISP出口信息，包括地址池和链接数
15 限制等。在转发流程中查询完路由表和策略路由后，再查询NAT策略表，获得出口ISP的信息，然后根据这些信息进行NAT操作。当出口ISP故障时，自动选择可用ISP，实现多个ISP之间的热备份。

具体的转发步骤如下所述：

- 1、根据报文的目的IP查询路由表，得到报文的转发出接口A；
- 20 2、根据系统配置信息，确定是否需要进行策略路由，如果不需要则使用源IP+出接口A查询NAT策略树，然后跳转到第5步；
- 3、根据复杂流分类的结果进行策略路由，得到报文的出接口B；
- 4、判断出接口B是否有效，如果有效则使用源IP+出接口B查询策略树，如果无效则使用源IP+出接口A查询NAT策略树，

5、根据NAT策略树查询结果选择出口的ISP，并对报文进行NAT操作，然后从选中的出口ISP链路中把报文发送出去；

多ISP出口备份通过两个手段实现：

1、对于策略路由，如果策略路由的出接口无效，则自动使用路由表的出接口。

2、对于路由表中的出接口，如果出接口无效，核心路由交换机的路由处理系统会自动进行路由重新计算，选择新的路由并下发到路由表中，从而实现多个ISP出口的备份。

10 为实现在多地址空间混合组网下，对多ISP出口的园区网设备的可控性要求，本发明“策略NAT”要实现以下三个关键功能点：

A 用户流量的出口不能仅由普通路由确定，必须完成完善的策略路由，同时实现普通路由对策略路由的备份。

15 B 必须支持同一个私网内的用户可以通过不同 ISP 提供的出口访问公网。并且同一私网用户从不同的出口访问公网时，该用户的私网地址应该转换成不同地址池中的公网地址，即用户地址空间与其出口空间不符时，就必须以其出口绑定的地址池进行 NAT 转换。

20 C 路由模块与 NAT 模块完全分离，路由模块（包括目的地址路由和策略路由）确定用户流量的出口，是否要做 NAT 以及选择哪个地址池，由 NAT 模块确定。

为了实现用户流量的出口不能仅由普通路由确定，必须完成完善的策略路由，同时实现普通路由对策略路由的备份的目的，利用普通路由对策略路由的备份，即在策略路由不可用时，用户流量自动依照
25 普通路由转发。本发明的实施例中以策略路由的下一跳查找路由表，

因为可用路由的下一跳通常是一个直联主机，能否命中 32 位掩码的直联主机路由，在本发明中做为策略路由是否可用的判断依据。如果命中 32 位掩码的直联主机路由，则说明此策略路由可用；否则，策略路由不生效，仍然以普通路由转发。

5

为了支持同一个私网内的用户可以通过不同 ISP 提供的出口访问公网，并且同一私网用户从不同的出口访问公网时，该用户的私网地址应该转换成不同地址池中的公网地址，即用户地址空间与其出口空间不符时，就必须以其出口绑定的地址池进行 NAT 转换。本发明的实施例实现了不同的用户流量由不同的 ISP 出口，以不同的地址池进行 NAT 转换。本发明的实施例中不再将地址池注册在全局模式下，而是与出接口绑定。同时为识别是否需要进行 NAT 转换和选择哪个 NAT 地址池，以出接口和源 IP 的组合做为关键字创建 NAT 策略树，记录地址池的绑定关系和分布式专用 NAT 转换板槽号。用户流量命中的路由表项中是否有公网标记做为搜索 NAT 策略树的使能开关。该标记由用户在与 ISP 对接的出接口上配置，路由凡落在这个出接口的，都会带上公网标记。根据出接口和源 IP 组成的联合为关键字，若能命中 NAT 策略树叶子，说明报文出口前需进行 NAT 转换，在 NAT 策略树叶子中得到地址池和 NAT 板号，报文转 NAT 板处理；否则若不能命中，说明此用户是公网地址，与此处接口的 ISP 属于同一编址空间，不需进行 NAT 转换，直接根据路由信息转下行用户板处理。

为了实现路由模块与 NAT 模块完全分离，路由模块（包括目的地址路由和策略路由）确定用户流量的出口，是否要做 NAT 以及选择哪

个地址池，由 NAT 模块确定。本发明的实施例采用路由模块与 NAT 模块完全分离保证两者在逻辑上划分清晰，功能上不再互相影响牵制。为实现不同用户各种复杂转发逻辑的组合，留下了充分的空间。

5 图 3 描绘了本发明的另一实施例的结合具体设备描绘的策略 NAT 转发流程图。

在步骤 210，利用目的 IP 查找路由表，根据路由表确定可能的下一跳。

在步骤 220，根据查找的路由表，判断是否存在多个下一跳，

10 如果在步骤 220 判断有多个下一跳，在步骤 230，这些多个下一跳进行流量分担；然后进入步骤 240，判断策略路由是否匹配成功；

如果在步骤 220 判断没有多个下一跳而仅有一个下一跳，直接进入步骤 240，判断策略路由是否匹配成功。

15 如果在步骤 240 判断策略路由匹配成功，进入步骤 250，利用策略路由的下一跳查路由表判断能否命中主机路由，如果在步骤 250 判断能命中主机路由，进入步骤 260，利用策略路由结果覆盖目的地址路由。然后，进入步骤 270，判断路由表项是否有公网标志。

20 如果在步骤 240 判断策略路由匹配不成功，及如果在步骤 250 判断不能命中主机路由，进入步骤 270。

如果在步骤 270 判断有公网标志，进入步骤 280，利用源 IP 和出口查 NAT 策略树，判断是否命中。

如果在步骤 280 判断命中，进入步骤 290，根据搜索结果，得到地址池号。然后，在步骤 310 经交换网，转分布式 NAT 处理设备，实现

NAT 转换。然后，在步骤 300，进入交换网。

如果在步骤 270 判断没有公网标志，及如果在步骤 280 判断未命中，在步骤 300，进入交换网。

最后，在步骤 320，按路由结果转发报文到出接口用户板。

5

以上所述，仅为本发明的实施例而已，其中所描述的方法都只是用做举例，非因此即局限本发明的权利范围，凡运用本发明说明书及附图内容的等效变化，均包含于本发明的权利要求范围内。

10

说明书附图

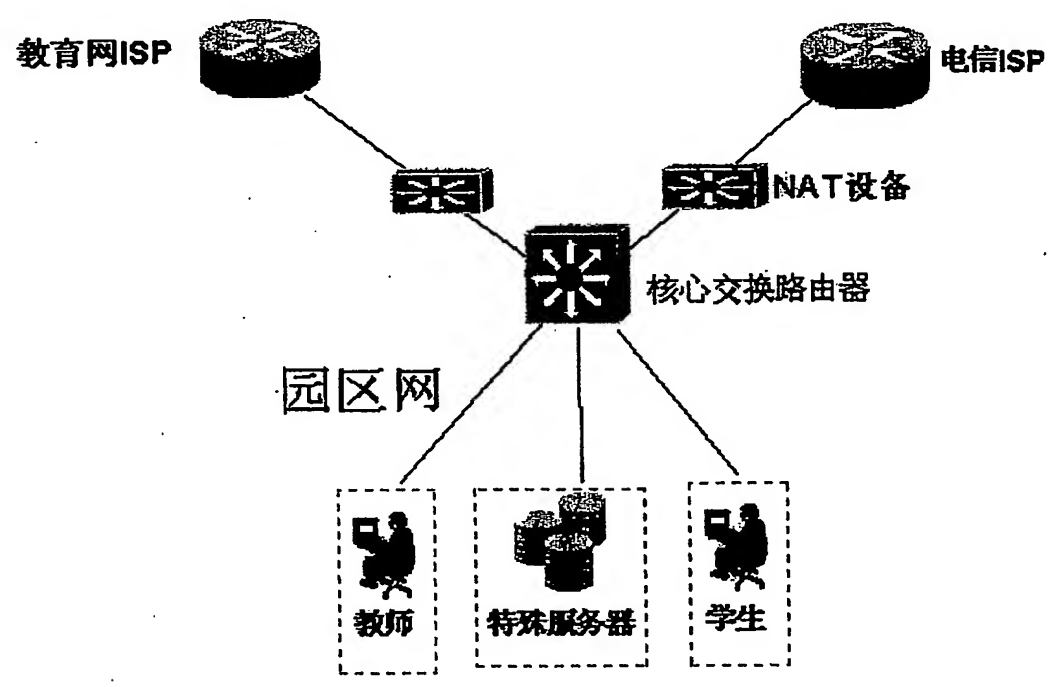


图 1

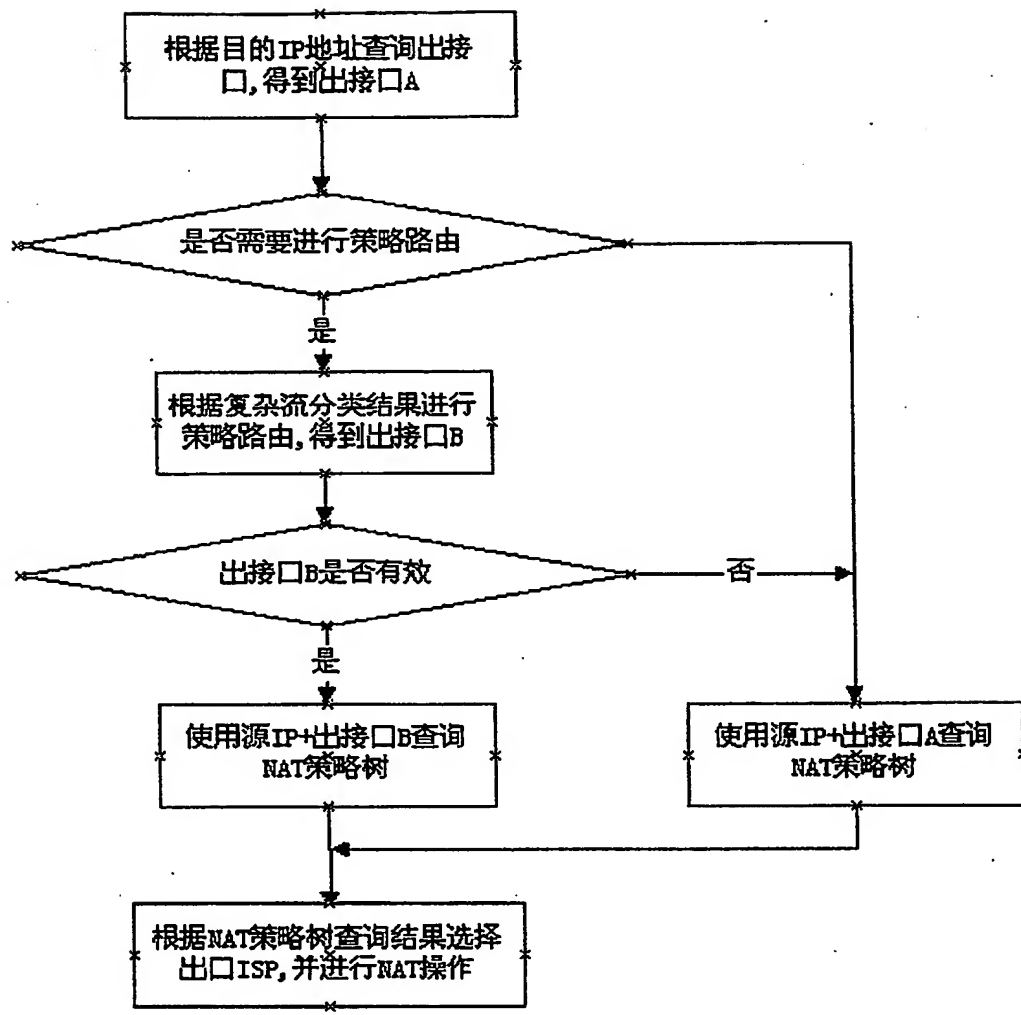


图 2

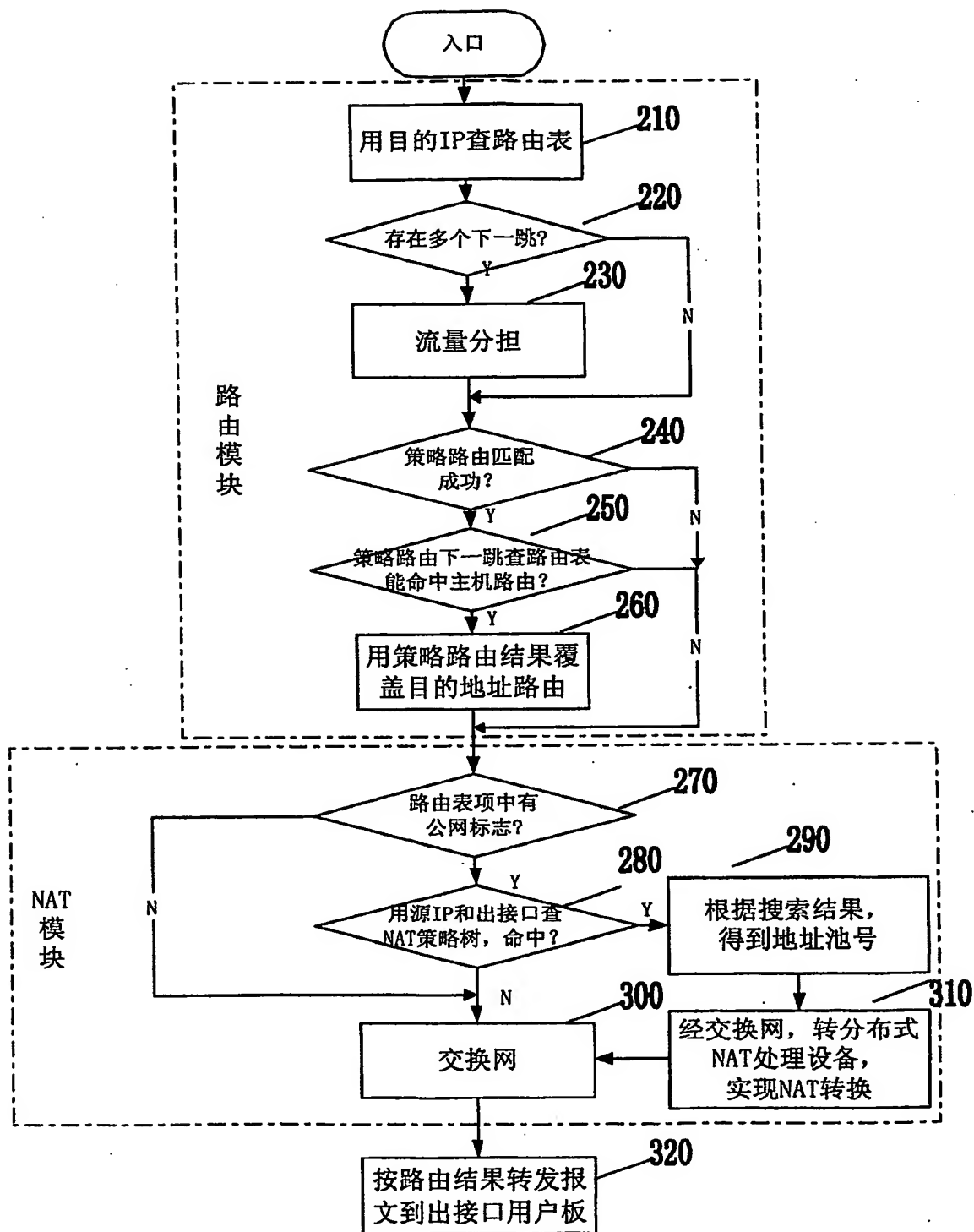


图 3

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/CN04/001456

International filing date: 15 December 2004 (15.12.2004)

Document type: Certified copy of priority document

Document details: Country/Office: CN
Number: 200310123979.3
Filing date: 19 December 2003 (19.12.2003)

Date of receipt at the International Bureau: 24 February 2005 (24.02.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record.

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

☒ **BLACK BORDERS**

☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**

☐ **FADED TEXT OR DRAWING**

☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**

☐ **SKEWED/SLANTED IMAGES**

☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**

☐ **GRAY SCALE DOCUMENTS**

☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**

☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**

☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.